Europ. J. Agronomy 82 (2017) 163–172

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

From grid to field: Assessing quality of gridded weather data for agricultural applications

Spyridon Mourtzinis [a,*], Juan I. Rattalino Edreira [b], Shawn P. Conley [a], Patricio Grassini [b]

[a] Department of Agronomy, University of Wisconsin-Madison, 1575 Linden Drive, Madison, WI 53706, USA
[b] Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583-0915, USA

ARTICLE INFO

Article history:
Received 2 August 2016
Received in revised form 19 October 2016
Accepted 22 October 2016
Available online 27 October 2016

Keywords:
Weather data
Simulation model
Crop yield
Phenology
Interpolation

ABSTRACT

High quality measured weather data (MWD) are not available in many agricultural regions across the globe. As a result, many studies that dealt with global climate change, land use, and food security scenarios and emerging agricultural decision support tools have relied on gridded weather data (GWD) to estimate crop phenology and crop yields. An issue is the agreement of GWD with MWD and the degree to which this agreement may influence the utility of GWD for agricultural research. The objectives of this study were: (i) to compare the agreement of two widely used gridded weather databases (GWDs) (Daymet and PRISM) and MWD, (ii) to evaluate their robustness at simulating maize growth and development, and (iii) to examine how GWD compare relative to weather data interpolated from existing meteorological stations for which MWD are available. The U.S. Corn Belt, a region that accounts for 43 and 34% of respective global maize and soybean production, was used as a case of study because of its dense weather station network and high-quality MWD. Historical daily MWD were retrieved from 45 locations across the region, resulting in *ca.* 1300 site-years. To test the accuracy of GWDs, separate simulations of maize yield and development were performed, separately for the two GWDs and MWD, using a well-validated maize crop model. For both GWDs, small biases were observed for temperature and growing degree-days in relation with MWD. However, accuracy was much lower for relative humidity, precipitation, reference evapotranspiration, and degree of seasonal water deficit. There was close agreement in duration of vegetative and reproductive phases between GWD and MWD, with root mean square error (RMSE) ranging from 3 to 7 days for the different crop phases and GWDs. However, robustness of GWDs to reproduce maize yields simulated using MWD was lower as indicated by the RMSE (18 and 24% of average yield for Daymet and PRISM, respectively). There was also a high proportion of site-years (20 and 32% for Daymet and PRISM, respectively) exhibiting a yield deviation >15% in relation to the yield simulated using MWD. Data interpolation using a dense weather station network resulted in lower RMSE% for simulated phenology and yields relative to GWDs. Findings from this study indicate that GWD cannot replace MWD as a basis for field-scale agricultural applications. While GWD appear to be robust for applications that only require temperature for prediction of crop stages, GWD should not be used for applications that depend on accurate estimation of crop water balance, crop growth, and yield. We propose that the evaluation performed in this study should be taken as a routinary activity for any research or agricultural decision tool that relies on GWD.

1. Introduction

Lack of measured daily weather data (MWD) at appropriate spatial resolution is a serious constraint to forecast current and future effect of weather on crop yields and to develop and use agricultural decision-support tools for crop and inputs management (Van Wart et al., 2013, 2015; Grassini et al., 2015). This phenomenon seems to be ubiquitous, even for important agricultural regions in developed countries such as the US Corn Belt (Fig. 1A). A fairly dense station network (2125 weather stations) has been established in this region (Fig. 1B). However, many of these stations are located at airports and cities and, therefore, MWD cannot reliably be used for agricultural applications. Likewise, most of these stations only measured rainfall and sometimes temperature but do not include other important variables for crop growth and yield such as solar radiation and humidity. When only weather stations

* Corresponding author.
E-mail address: mourtzinis@wisc.edu (S. Mourtzinis).

http://dx.doi.org/10.1016/j.eja.2016.10.013

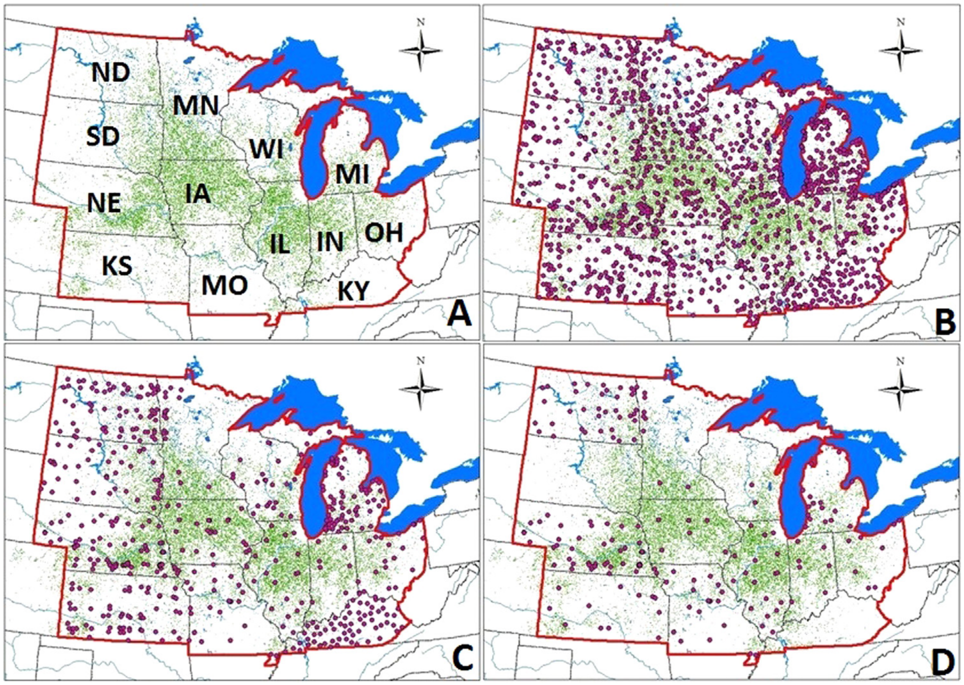

Fig. 1. (A) Map of the US Corn Belt, including portions of the Central Great Plains region (ND-North Dakota, SD-South Dakota, NE-Nebraska, KS-Kansas, MN-Minnesota, IA-Iowa, MO-Missouri, WI-Wisconsin, IL-Illinois, MI-Michigan, IN-Indiana, OH-Ohio, and KY-Kentucky). Red line shows the extent of the region. Area sown with maize and soybean is shown in green. (B) Location of all active and inactive meteorological stations collecting daily weather data. (C) Distribution of stations located in agricultural areas and collecting data for all agronomically relevant variables. (D) Distribution of active weather stations with publicly available, long-term (>15 years) daily weather data records for all agronomically relevant weather variables and located in agricultural areas. Each circle indicates the location of a meteorological station.

measuring all weather variables needed for agricultural applications (radiation, temperature, humidity, and wind speed) located in agricultural areas are considered, total number reaches 529 (including inactive stations) and average coverage reaches 4026 km^2 per station (Fig. 1C). Additionally, the distribution of the stations is uneven resulting in states with a dense network (*e.g.,* Kentucky, North Dakota, and Nebraska) and states with a sparse to almost non-existing network (*e.g.,* Minnesota and Wisconsin). Many stations have stopped to be operated in recent years or data are not publicly available (*e.g.,* South Dakota). Other stations have started to be operated in recent years *e.g.,* Kentucky; these stations are of limited used for agricultural applications that require long-term records to account for weather variability. Considering only active stations with complete long-term daily data (>15 years and missing observations <10%), the number for which is possible to obtain reliable, long-term daily weather data is only 184, which translates to an average coverage of 11,575 km^2 per station (Fig. 1D).

The current trend of increasingly higher volume of gridded weather data (GWD), in contrast to increasing scarcity of (or lack of access to) MWD, has led researchers to use GWD as basis for assessments on climate change, food security and land use (Mourtzinis et al., 2015, 2016; Overpeck et al., 2011; Van Wart et al., 2013 and references cited therein). Likewise, GWD have started to be used for agricultural-related research and decision-support tools to guide crop and input management (*e.g.,* https://www.ral.ucar.edu/solutions/decision-support-tools-farmers, https://ifdc.org/decision-support-tools/; Daly et al., 2012; Miner et al., 2013). GWD are typically generated from satellite images or interpolations from meteorological stations using a collection of tortuous and empirical algorithms to produce gridded estimates of daily weather parameters at the desired spatial and scale. Influence of the level of weather data spatial aggregation on crop simulations has been investigated in previous studies (Angulo et al., 2013; Zhao et al., 2014; Rezaei et al., 2015). However, very few studies have evaluated these GWDs on their agreement with MWD from stations located within the same grid have indicated lack of agreement and important biases (Ramirez-Villegas and Challinor, 2012; Van Wart et al., 2013, 2015). However, these previous evaluations have been based on GWDs with coarse spatial resolution (from 3000 to 70,000 km^2), such as NASA-POWER National Aeronautics & Space Administration; http://power.larc.nasa.gov/), NCEP (National Center for Environmental Prediction/Department of Energy; http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html), and CRU (Climate Research Unit; http://badc.nerc.ac.uk/data/cru/). There is lack of a robust assessment of most recent GWDs, with more granular spatial resolution (<20 km^2), relative to their potential use for agricultural applications. And while some of these GWDs have been evaluated by comparing them against MWD, these previous evaluations can benefit from using a crop simulation model that integrates the effects of weather, management practices, soils, and crop cultivars on crop development, growth, and yield (Van Ittersum et al., 2013; Van Wart et al., 2013, 2015).

In the present study, we evaluated accuracy of state-of-art GWD containing daily weather data at high level of spatial resolution. We used the US Corn Belt as a case of study because it covers one of the most important agricultural areas of the world, with substantial spatial weather and soil variation, and has a fairly dense network of weather stations with long-term daily MWD including all the weather variables needed for agricultural applications. We assessed accuracy of GWD by comparing their agreement with high-quality MWD as well as associated simulated crop phenology and yields based on a well-validated crop model and using actual dominant soil and management practices for each location. As an alternative approach to GWD, we also evaluated the accuracy of weather data interpolated from nearby weather stations and how this accuracy depended upon the density of the weather station network.

2. Materials and methods

2.1. Measured and gridded weather data sources

MWD were retrieved from 45 locations across the U.S Corn Belt (Fig. 2, see Supplementary Table S1 in the online version at DOI: 10.1016/j.eja.2016.10.013). These stations belong to the MESONET weather network (http://mrcc.isws.illinois.edu/gismaps/mesonets.htm). Available years of MWD ranged between 12 and 35 across locations, resulting in *ca.* 1300 site-years of observations included in the study. Measured variables include incident solar radiation, maximum and minimum air temperatures (Tmax and Tmin, respectively), relative humidity (RH), wind speed, and precipitation. Reference evapotranspiration (ETo) was calculated using the grass-referenced FAO–Penman–Monteith equation (Allen et al., 1998). Station distribution encompassed a wide range of weather conditions across the US Corn Belt. For example, annual precipitation ranged from 800 mm (Illinois) to 350 mm (western Nebraska). Further details on weather data sources and quality control measures can be found in Morell et al. (2016). MWD were used here as a benchmark against which GWD were evaluated to evaluate their accuracy.

Two publicly-available, "state-of-art" GWDs were used in the present study: Daymet (Daily Surface Weather Data on a 1-km Grid for North America; Thornton et al. (2014)) and PRISM (Parameter-elevation Relationships on Independent Slopes Model; PRISM Climate Group (2004)). Weather data in these databases are reported on a daily basis and at high spatial resolutions: 1 km^2 (Daymet) and 16 km^2 (PRISM). Daymet includes data for North America and Hawaii; including Canada, Mexico, the US, Bermuda, and Puerto Rico whereas PRISM provides data for the conterminous US. The algorithms used to generate Daymet and PRISM have been described in detail elsewhere (Thornton et al., 1997; Daly et al., 2008). These two databases represent the two current most popular sources of daily weather data in USA and they have served as basis for several studies about the effect of climate change in agriculture, some of them published in high-impact journals (Lobell et al., 2014; Miner et al., 2013; Prein et al., 2016). GWD were retrieved for those grids that coincided with the location of the 45 weather stations shown in Fig. 2. Daymet and PRISM included the same weather variables as those available for the MWD collected from the 45 locations. An exception was solar radiation (PRISM) and RH (both Daymet and PRISM). The missing variables were retrieved from best available data sources or estimated with standard equations, consistent with most typical approaches followed in previous studies to estimate these variables in absence of MWD (*e.g.,* Lobell et al., 2014). The National Aeronautics and Space Administration's POWER database (NASA-POWER; 12,000 km^2 resolution) was used as a source of solar radiation for PRISM. NASA-POWER solar radiation has shown to be accurate when compared against measured solar radiation data for agricultural regions with flat terrain, as it is the case of the US Corn Belt (Van Wart et al., 2013, 2015; Bai et al., 2010; White et al., 2011). RH was estimated from GWD vapor pressure data (Daymet) or from Tmax and Tmin (PRISM) following the procedures described by Allen et al. (1998). For both GWDs, ETo was calculated using the grass-referenced FAO–Penman–Monteith equation (Allen et al., 1998).

2.2. Simulation of maize development and yield

The integrative impact of weather variables on crop development, growth, and final yield was evaluated using a mechanistic crop simulation model. Yield potential was simulated for irrigated and rainfed maize using Hybrid-Maize simulation model (Yang et al., 2004, 2006). Hybrid-Maize is a crop simulation model that simulates maize growth and development on a daily time step, for both rainfed and irrigated crops, as influenced by weather (e.g., temperature, solar radiation, precipitation), soil properties (e.g., soil depth, texture), and key management practices such as cultivar maturity, sowing date, and plant density (Yang et al., 2004; http://hybridmaize.unl.edu). The model has been calibrated with data collected from well-managed rainfed and irrigated crops that grew without nutrient limitations and biotic stresses (Yang et al., 2004). This model has been rigorously tested on its ability to reproduce measured yields across a wide range of environments and managements (Grassini et al., 2009).

Rainfed and irrigated maize production prevailed (>90% of maize planted area) in respective 29 and 4 of the 45 locations; hence, simulations were performed only for rainfed or irrigated maize at these sites (Fig. 2). Both water regimes were simulated for those locations (12 out of 45) where both irrigated and rainfed crop production are important (Fig. 2). For irrigated crops, yield potential was estimated under the assumption of optimum management, that is, well-adapted crop cultivars grown in absence of yield limiting and reducing factors such as water and nutrient deficiencies and biotic stresses (Evans, 1993). Hence, yield potential is determined by solar radiation, temperature, management practices such as sowing date and plant density, and crop traits that govern crop season length and capture and conversion of solar radiation into crop biomass (Van Ittersum et al., 2013). For rainfed crops, simulations of yield potential also accounted for the influence of water supply amount and distribution and soil and terrain properties related with water availability (soil depth, water holding capacity, slope). Hereafter, simulated yield potential for irrigated and rainfed crops is referred as simulated yield.

Separate simulations were performed based on three sources of weather data (measured, Daymet, and PRISM). Dominant soil type and management practices (sowing date, hybrid maturity, and plant density) at each location were used as basis for the simulations (Fig. 2, see Supplementary Table S1 in the online version at DOI: 10.1016/j.eja.2016.10.013). Details on soils and management data sources can be found in Morell et al. (2016). For each site-year, soil water content at sowing was determined dynamically by initializing the model run at harvest of the prior crop (*i.e.*, about 6 months before sowing), assuming 50% of available soil water content at that time and actual weather data from that prior harvest to sowing. The same set of soil and management input data for each location was used consistently across simulations based on the three different sources of weather data. Simulated crop parameters included days from emergence to silking (vegetative days), days from silking to physiological maturity (reproductive days), aboveground dry matter, harvest index, and grain yield (reported at standard grain moisture content of 15.5%).

2.3. Comparison of measured and gridded weather data sources

Average solar radiation, Tmax, Tmin, RH, total precipitation, and total ETo were calculated for the period between April 1st and September 30th for each site-year, separately for the MWD and the two GWDs. In most of the western Corn Belt, precipitation is not adequate to replace water losses due to evapotranspiration (Grassini et al., 2009). Given the importance of water balance on crop yield, we calculated water deficit as the difference between total ETo and total precipitation for each site-year. Growing degree-days (GDD, °Cd) were calculated as the sum of mean temperature from sowing to simulated physiological maturity after subtracting a base temperature of 10 °C and using an upper 30 °C cutoff (Ritchie and Hanway, 1982).

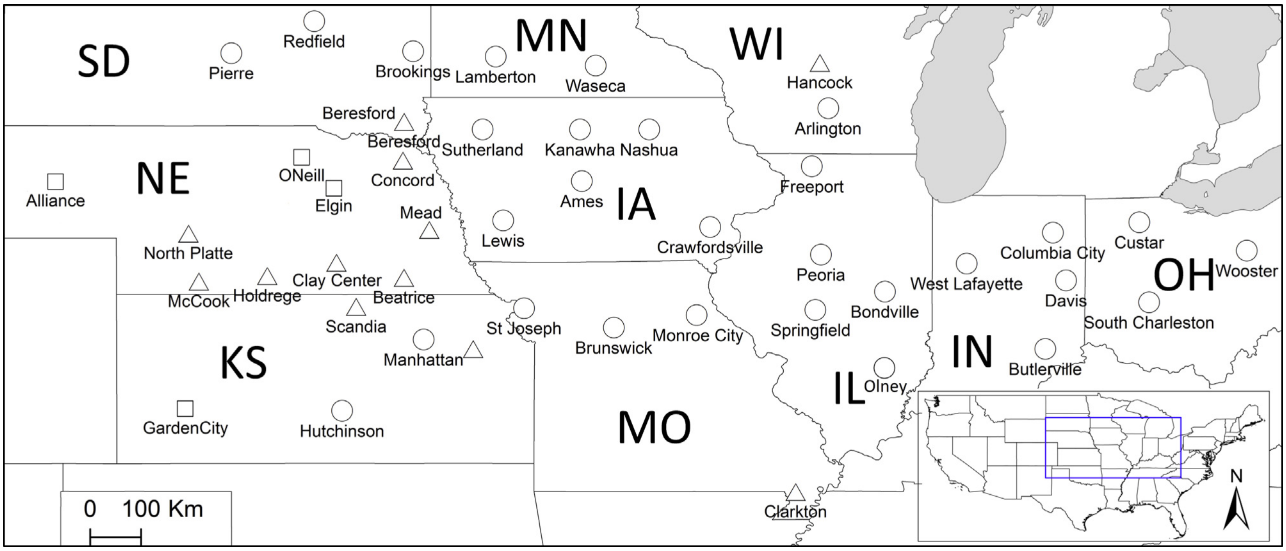

Fig. 2. Distribution of weather stations used for this study. For each weather station, maize yield potential was simulated for rainfed (circles), irrigated (squares) or both water regimes (triangles) depending upon the prevalence of each water regime at each site. The inset shows the target region within US (SD-South Dakota, NE-Nebraska, KS-Kansas, MN-Minnesota, IA-Iowa, MO-Missouri, WI-Wisconsin, IL-Illinois, MI-Michigan, IN-Indiana, and OH-Ohio).

For a given weather or crop parameter, agreement and biases between GWDs and MWD were assessed with the absolute mean error (ME) and root mean square error (RMSE):

$$ME = \frac{\sum_{i=1}^{n} (yiM - yiG)}{n} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (yiM - yiG)^2}{n}} \quad (2)$$

where y_i^M is the variable derived from the MWD for the ith site-year, whereas y_i^G is the variable-specific GWD for the ith site-year. The ME measures the magnitude and sign of the bias, while RMSE provides a measure of the degree of agreement between weather data sources. RMSE was also calculated as percentage (RMSE%) of the measured mean for a given weather variable or the mean simulated yield based on MWD. Finally, linear regression analysis was performed to detect biases in the relationship between GWD and MWD and the coefficient of determination (r^2) was also calculated.

2.4. Investigating weather data interpolation as an alternative to GWD

We explored the possibility of using weather data interpolated from MWD collected at meteorological stations as an alternative to GWD. At issue is how dense a weather network needs to be so that resulting interpolation can be considered robust. To examine the effect of weather station density on accuracy of the interpolated weather data, Nebraska (NE) and Illinois (IL) were used as case studies (Fig. 3). NE has a dense MESONET network with 63 active stations distributed across a total area of 200,520 km^2 (3180 km^2 per station). In contrast, IL has a sparse weather station network (18 active stations) distributed across an area of 149,932 km^2 (8330 km^2 per station). A few stations from adjacent states were also included in the analysis to increase the spatial coverage and density of the state weather networks.

The method followed to interpolate daily weather data uses the coordinates of the target site to calculate the distance of the three nearest stations from the target site (Yang and Torrion, 2013; http://hybridmaize.unl.edu/weather-interpolator). Subsequently, it applies inverse distance weighting to estimate a daily value for a given weather variable based on the MWD at the three stations. The inverse distance weighting method calculates a weighted average of the proximate known observations with weights being a decreasing function of distance from the target point (Tang et al., 1996). In the present study, we interpolated weather data for a subset of the locations shown in Fig. 2, hereafter called 'target sites', which included 7 and 5 sites in NE and IL, respectively (Fig. 3A and B, black stars). To determine how density of weather stations influenced the accuracy of interpolated weather data, separate interpolations were performed for five different network density scenarios: 100, 75, 50, 25, and 5% of all stations within a state. For example, in the 100% density scenario, all available active stations were used to interpolate the weather data to the target sites. For the 75% density scenario, 25% of stations were removed prior to the interpolation. To ensure consistency, for all density scenarios and states, individual stations were removed sequentially, starting from the closest one to the target site, until the desired density was achieved. The 5% density scenario was not assessed for IL because it would have led to an extremely low number of weather stations (only one). Maize yield and phenology were simulated for the target sites using all available years of interpolated weather data. For a given state-density scenario, all site-year simulations were pooled and compared against simulations based on MWD. Independent RMSE values were-calculated for each state-density scenario. Only rainfed maize was considered for this assessment.

3. Results and discussion

3.1. Comparison of seasonal weather variables

Comparison of average Tmax and Tmin indicates reasonably good agreement between GWD and MWD, with RMSE consistently representing ≤ 5% of the mean based on the MWD (Fig. 4A–D). Despite good agreement between GWD and MWD solar radiation (RMSE% = 8), correlation was poor due to the narrow range of average solar radiation across site-years (Fig. 4E and F). In all three cases (Tmax, Tmin, and solar radiation), the majority of data points (>98%) fall between ±15% of the measured values, and this was consistent for both GWDs. In contrast, RH estimated for both Daymet and PRISM exhibited poor agreement with measured RH, with respective RMSE% of 18 and 13% (Fig. 4G and H). On the one hand, lack of agreement between Daymet and measured RH was associated

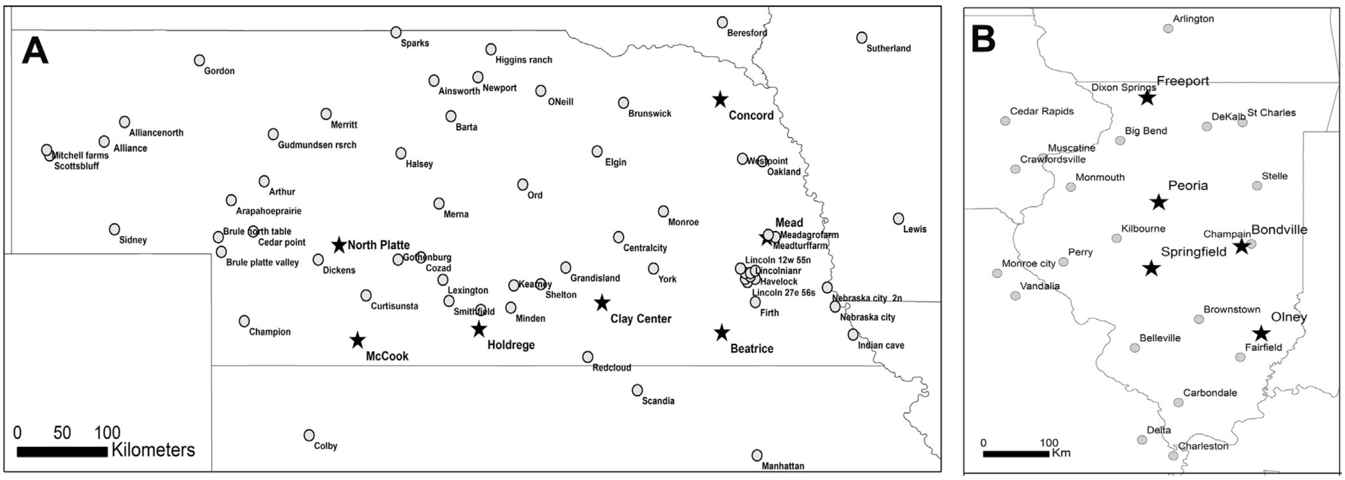

Fig. 3. Distribution of weather stations across Nebraska (A) and Illinois (B). Measured weather data (MWD) were interpolated from the three stations located near each target site (black stars) and, subsequently, simulations based on interpolated weather were evaluated against simulations using MWD collected by a meterological station located at the target site.

with the poor agreement for actual vapor pressure between data sources (RMSE% = 24, See Supplementary Fig. S1a in the online version at DOI: 10.1016/j.eja.2016.10.013). On the other hand, given the close agreement between PRISM and measured Tmax and Tmin, the disagreement in RH between these two data sources seems to be associated with uncertainties in the method used to estimate RH and actual vapor pressure from Tmax and Tmin (Fig. 4, see Supplementary Fig. 1b in the online version at DOI: 10.1016/j.eja.2016.10.013).

Given the close agreement between measured and GWD Tmax and Tmin, it was not surprising that GDD calculated based on GWDs were in close agreement (RMSE% = 4) with GDD estimated based on MWD (Fig. 5A and B). In contrast, agreement was lower for total precipitation and ETo, with RMSE% ranging from 10 to 29% across parameters and GWDs (Fig. 5C–F). In addition, there was a consistent tendency of GWD to overestimate total precipitation (ME: 59 and 41 mm for Daymet and PRISM, respectively). Likewise, there was a remarkable bias between MWD and PRISM for ETo (ME = 253 mm). In the case of Daymet, the bias was less evident (ME = -4 mm) although there was a clear tendency to under- and over-estimate ETo in the high and low range of MWD (Fig. 5E and F). Poor agreement between measured and GWD ETo was related to the poor agreement for RH observed among MWD and GWDs (Fig. 4G and H). These biases between MWD and GWD explained the underestimation of the actual degree of water deficit by the two GWDs, especially in the case of PRISM (Fig. 5G and H). To summarize, while the GWD appeared robust at reproducing the measured GDD, they performed poorly at reproducing the actual degree of water deficit across site-years.

3.2. *Comparison of simulated yields based on measured and gridded weather data*

Separate simulations of maize phenology and final yield were performed using the three sources of weather data to evaluate the ability of GWD to reproduce the results based on MWD. The effect of using different weather data sources for simulating length of vegetative (emergence-silking) and reproductive phases (silking-physiological maturity), and grain yield is shown in Fig. 6. There was a close agreement for duration of simulated vegetative and reproductive phases between GWDs and MWD (RMSE% ≤ 12) (Fig. 6A–D). Similarly, there was no bias for vegetative and reproductive days between GWD and MWD as indicated by the associated low MEs (<1 d) and very high correlation (r^2 = 0.87). In 99% of the cases, simulations using GWDs were within ±15% from the simulated value based on MWD and these findings were consistent across individual locations (see Supplementary Figs. S2 and S3 in the online version at DOI: 10.1016/j.eja.2016.10.013). These results are consistent with the close agreement found for Tmax, Tmin, and GDD between GWD and MWD (Figs. 4 and 5). However, agreement between GWD and MWD was clearly weaker for duration of the reproductive phase compared with the vegetative phase (RMSE% = 12 and 4, respectively) (Fig. 6A–D). While duration of the vegetative phase depends only on temperature, length of the reproductive phase can be shortened due to severe water stress and frost occurrence. Indeed, the discrepancy in RMSE% calculated for the vegetative and reproductive phases is due to (i) weaker agreement between GWD and MWD for water deficit than for temperature (Fig. 5) and (ii) slightly lower frequency of site-years with frost occurrence in simulations based on GWD relative to MWD (26, 26, and 36% for Daymet, PRISM and MWS, respectively).

In contrast to phenology, simulated yields using GWD exhibited poorer agreement relative to yields simulated using MWD, with RMSE% of 18% (Daymet) and 24% (PRISM), and much lower correlation (r^2 < 0.67) (Fig. 6E and F). There was also a consistent tendency of PRISM to overestimate yields (ME = −1.4 Mg ha^{-1}). Based on RMSE% calculated for each location-water regime case, Daymet outperformed PRISM on the ability to reproduce simulated yields based on MWD simulated yields in 63% of the cases while PRISM outperformed Daymet in 27% of the cases (see Supplementary Fig. S5A and B in the online version at DOI: 10.1016/j.eja.2016.10.013). In the remaining cases (10%), RMSE% was similar for PRISM and Daymet. The contrasting performance can be attributed mainly to the better accuracy of Daymet to reproduce the actual water deficit (RMSE%: 29 *versus* 68%) and ETo (RMSE%: 10 *versus* 29%) compared with PRISM. Similar results were found for other simulated variables, such as aboveground dry matter and harvest index (Supplementary Fig. S4A–D in the online version at DOI: 10.1016/j.eja.2016.10.013). Daymet performed substantially better than PRISM in simulating both aboveground dry matter (RMSE%: 12 *versus* 16) and harvest index (RMSE%: 15 *versus* 18). PRISM grossly underestimated aboveground dry matter (ME: −2 Mg ha^{-1}), especially in harsh, low-yield rainfed environments.

Despite the better performance of Daymet compared with PRISM, agricultural applications that rely on Daymet for estimation of seasonal water deficit, crop growth, and yield are still subjected to a high degree of uncertainty. While Daymet simulated yields did not exhibit a consistent bias in relation with simulated yields

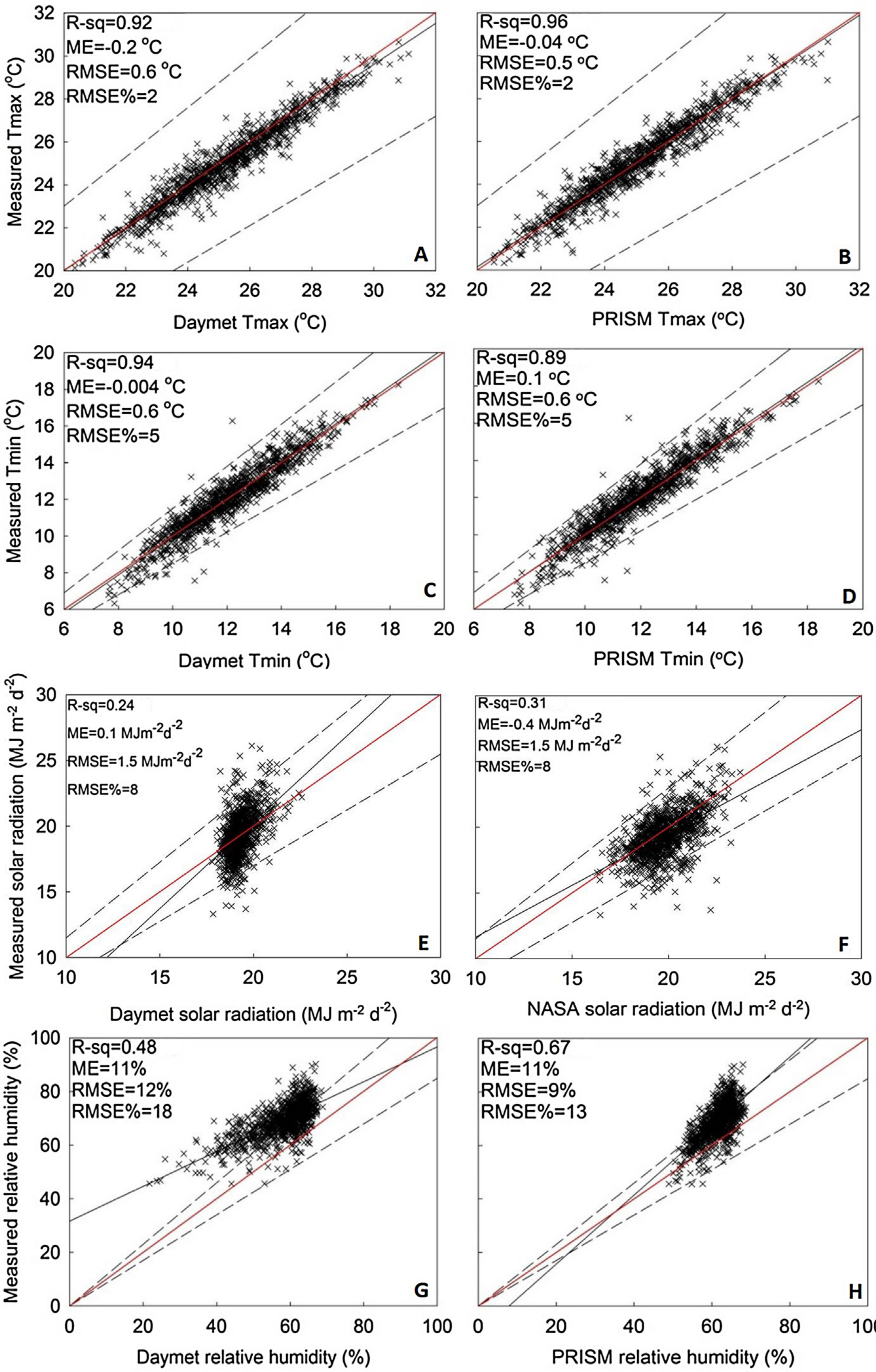

Fig. 4. Comparison between measured (MWD) and gridded weather data (GWD) for average Tmax (A and B), Tmin (C and D), solar radiation (E and F), and relative humidity (G and H) for two sources of GWD: Daymet (left) and PRISM (right). Each datapoint indicates the mean value for the April 1st–September 30th time period, which coincides roughly with the maize crop season in the US Corn Belt. Red line indicates y = x, while dashed lines show ±15% deviation from the y = x line. Solid black line is the fitted linear regression. Coefficient of determination (R-sq), mean error (ME), root mean square error (RMSE), and RMSE as% of the mean MWD (RMSE%) are also shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

based on MWD, there was a large number of cases (20%) that were beyond ±15% (Fig. 6E). Furthermore, detailed analysis across locations indicated that agreement and correlation between Daymet yields and simulated yields based on MWD changed erratically across sites-years and water regimes (see Supplementary Fig. S5A and B in the online version at DOI: 10.1016/j.eja.2016.10.013). For example, Daymet rainfed simulated yields were in close agreement with yields based on MWD at Clarkton MO, but, in contrast, there was a high disagreement at Beresford SD (RMSE%: 12 *vs.* 35, respectively). This pattern was not associated with spatial variation in weather variables, terrain attributes, or weather network.

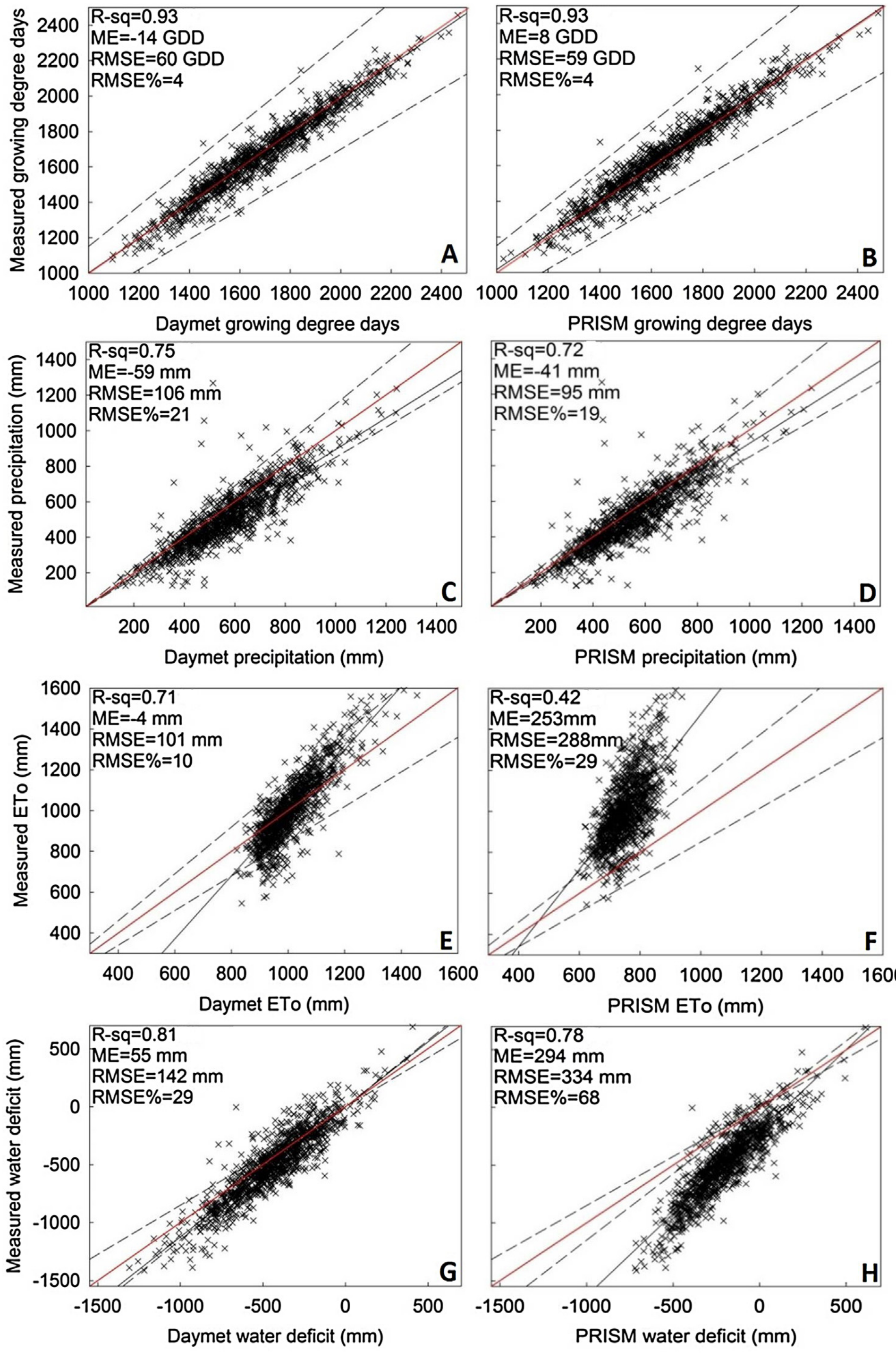

Fig. 5. Comparison between measured (MWD) and gridded weather data (GWD) for total growing degree-days (A and B), precipitation (C and D), reference evapotranspiration (E and F), and water deficit (G and H) for two sources of GWD: Daymet (left) and PRISM (right). Each datapoint indicates the mean value for the April 1st–September 30th time period, which coincides roughly with the maize crop season in the US Corn Belt. Red line indicates y = x while dashed lines show ±15% deviation from the y = x line. Solid black line is the fitted linear regression. Coefficient of determination (R-sq), mean error (ME), root mean square error (RMSE), and RMSE as% of the mean MWD (RMSE%) are also shown.

Despite the better performance of Daymet compared with PRISM, agricultural applications that rely on Daymet for estimation of seasonal water deficit, crop growth, and yield are still subjected to a high degree of uncertainty. While Daymet simulated yields did not exhibit a consistent bias in relation with simulated yields based on MWD, there was a large number of cases (20%) that were beyond ±15% (Fig. 6E). Furthermore, detailed analysis across locations indicated that agreement and correlation between Daymet yields and simulated yields based on MWD changed erratically across sites-years and water regimes (Supplementary Fig. S5A and B). For example, Daymet rainfed simulated yields were in close agreement with yields based on MWD at Clarkton MO, but, in contrast, there was a high disagreement at Beresford SD (RMSE%: 12 *vs.* 35, respectively). This pattern was not associated with spatial variation in weather variables, terrain attributes, or weather network.

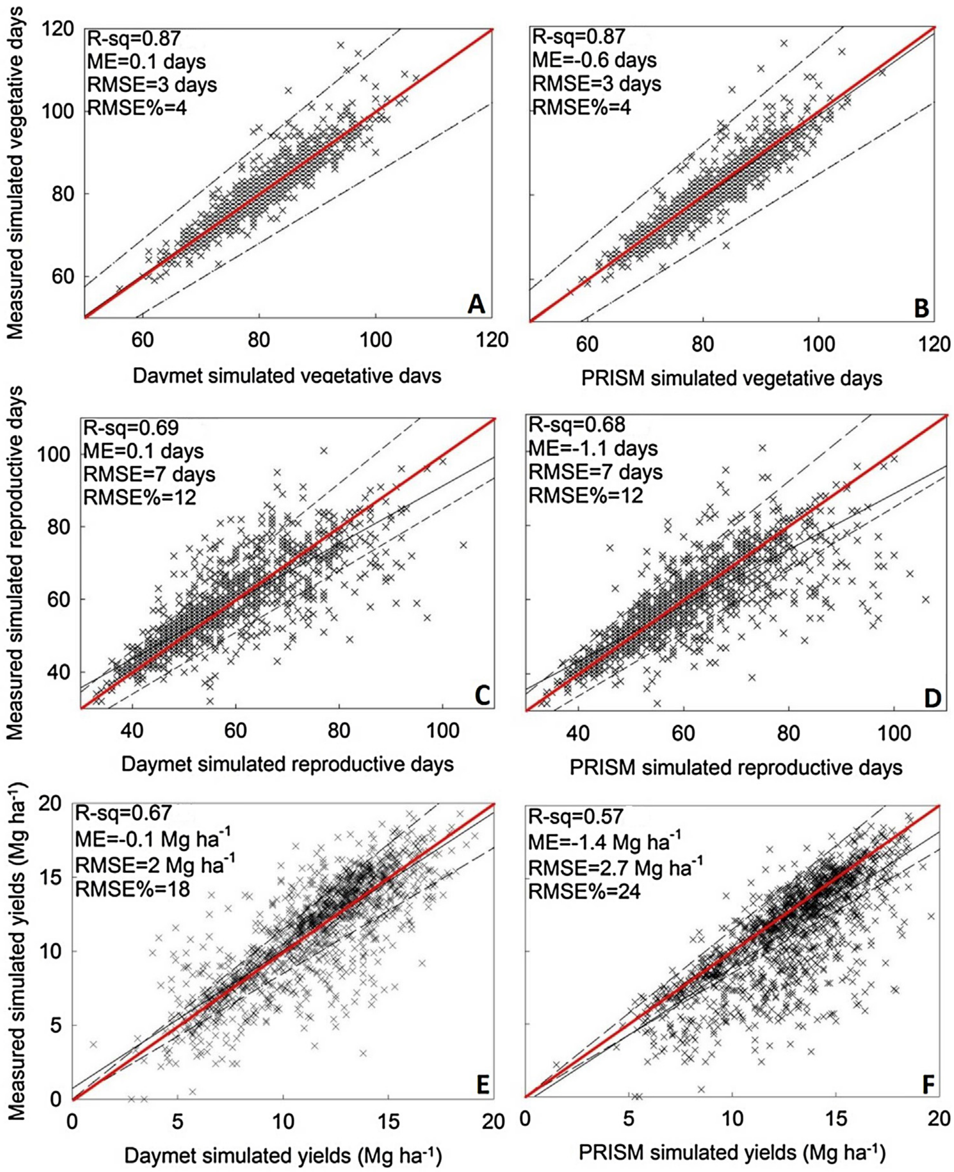

Fig. 6. Comparison between measured (MWD) and gridded weather data (GWD) for simulated vegetative (emergence to silking) days (A and B), reproductive (silking to physiological maturity) days (C and D), and grain yield (E and F) for two sources of gridded weather data: Daymet (left) and PRISM (right). Red line indicates y = x while dashed lines show ±15% deviation from the y = x line. Solid black line is the fitted linear regression. Coefficient of determination (R-sq), mean error (ME), root mean square error (RMSE), and RMSE as% of the mean based on MWD (RMSE%) are also shown.

3.3. Influence of weather station network density on accuracy of interpolated weather data

We tested an alternative approach to generate site-specific weather data based on triangulating MWD from nearest weather stations. Our findings indicate that the quality of the interpolated weather data was highly dependent upon the density of the underpinning weather network (Fig. 7). As expected, the denser the station network, the greater the agreement between simulations based on interpolated *versus* MWD-based simulations. In almost all cases, RMSE values for NE were lower that for IL given the same weather station coverage. This finding was unexpected given the greater spatial variation in weather, in particular precipitation, in NE compared with IL. Interpolating weather data using ≥75% of all available stations in NE (equivalent to < 4260 km^2 per station) resulted in simulated yields or phenology that had similar or even closer agreement with simulations based on MWD relative to simulations using Daymet (RMSE% < 20%) (Fig. 7C). In contrast, area coverage per station in IL was > 8300 km^2 in all cases, resulting in higher RMSEs relative to Daymet across all weather station density scenarios. In other words, the low station density in IL did not allow to generate interpolated weather data with superior performance over GWDs to reproduce the MWD simulated yields or phenology. Hence, our findings indicate that quality of interpolated weather data will ultimately depend on the density and distribution of the underpinning weather station network that was used as basis for the interpolation.

4. Conclusions

Results from this study showed that accuracy of the weather data is a key factor for reliable crop modeling and decision-support tools that require daily weather data. This study expands past assessment of coarse-resolution GWD (Van Wart et al., 2013, 2015)

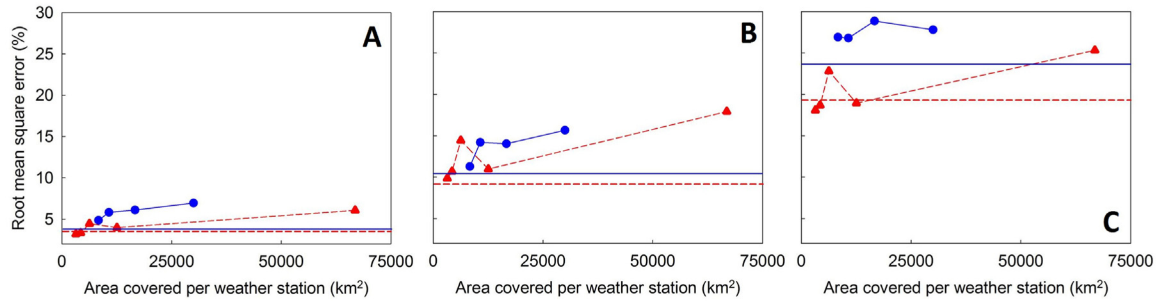

Fig. 7. Agreement in vegetative days (A), reproductive days (B), and grain yield (C) between simulations based on interpolated and measured weather data (MWD) as influenced by weather station density in Illinois (blue circles and lines) and Nebraska (red triangles and lines). Agreement was quantified with the root mean square error, expressed as a percentage of the mean based on MWD (Root mean square error%-RMSE%); each RMSE% value was calculated based on all site-years simulated for a given state-density combination. Horizontal lines indicate RMSE% calculated based on agreement between Daymet *versus* MWD simulations for each state. See Material and Methods for a detailed description of the methodology.

to more recent high-resolution GWDs released for U.S. Despite the improvements in accuracy that we found, on particular for temperature and simulated crop phenology, there are still important sources of uncertainty such as water deficit and simulated yields. This implies that outcomes from studies or decision-support tools based on GWD to assess impact of climate on crops have a high degree of uncertainty. While, on average, simulated yields using Daymet were in closer agreement with simulations based on MWD compared with PRISM (RMSE%: 18 *vs.* 24%), there were still large differences in simulated yield in 20% of the site-years. And, perhaps more importantly, these differences were not predictable as they were not associated with any spatial pattern in weather, topography, or weather network. While GWD might be useful for applications that only require temperature, such as crop stage prediction or quantification of early frost risk, water deficit and simulated yields for specific site-years are highly uncertain and there is no way to *ex-ante* predict the magnitude and direction of the bias, which undermines utility of GWD for field-specific or real-time agronomic applications. Nevertheless, in occasions that limited amount of MWD are available, the GWD can be calibrated based on their correlation with the existing MWD, which could ultimately reduce bias and improve their quality (*e.g.*, Van Wart et al., 2015).

The two GWD evaluated in the present study were created for the U.S., where high-quality MWD are available at relatively high spatial resolution. Hence, if MWD were used as foundation to develop these GWDs, one would have expected *a priori* good agreement between MWD and GWD; however, this was not the case for many of the weather variables tested here and corresponding simulated yields. Therefore, uncertainty associated with use of GWDs is likely to increase in regions of the world where MWD are less available or simply do not exist.

The findings reported here also highlight the importance of weather station network density in agricultural regions. Indeed, we hypothesize that no improvement in the methodology to create GWD can compensate for lower number of weather stations as it has been the trends in USA and many other countries. Hence, we emphasize the need to maintain and expand weather station networks in agricultural regions across the globe, which, in turn, can help generate interpolated weather data with improved accuracy in relation with existing GWD. Given the growing popularity and use of GWD for agricultural applications, and the increasing number of GWD that exist or are being created (Overpeck et al., 2011), we propose that the evaluation performed in this study should be taken as a regular activity for any kind of research or agricultural application that rely on GWD.

Acknowledgements

We acknowledged the North Central Soybean Research Program (NCSRP), Nebraska Soybean Board, and Wisconsin Soybean Marketing Board for their support to this project. We also thank Drs Haishun Yang (University of Nebraska-Lincoln) and Jessica Torrion (Montana State University) for their help at using the weather interpolator, DuPont Pioneer® agronomists for helping to retrieve plant maturity and density information, the USDA-Risk Management Agency (RMA) for providing sowing date data, and Dr Francisco Morell for helping collect the weather and soil data used for the simulations.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration: guidelines for computing crop water requirements. In: FAO Irrigation and Drainage Paper 56. FAO, Rome, Italy.
Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P., Ewert, F., 2013. Characteristic 'fingerprints' of crop model responses to weather input data at different spatial resolutions. Eur. J. Agron. 49, 104–114.
Bai, J., Chen, X., Dobermann, A., Yang, H., Cassman, K.G., Zhang, F., 2010. Evaluation of NASA satellite- and model derived weather data for simulation of maize yield potential in China. Agron. J. 102, 9–16.
Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol. 15, 2031–2064, http://dx.doi.org/10.1002/joc1688.
Daly, C., Widrlechner, M.P., Halbleib, M.D., Smith, J.I., Gibson, W.P., 2012. Development of a new USDA plant hardiness zone map for the United States. J. Appl. Meteorol. 51, 242–264, http://dx.doi.org/10.1175/2010JAMC2536.1.
Evans, L.T., 1993. Crop Evolution, Adaptation and Yield. Cambridge University Press, Cambridge.
Grassini, P., Yang, H.S., Cassman, K.G., 2009. Limits to maize productivity in Western Corn-Belt: a simulation analysis for fully irrigated and rainfed conditions. Agric. For. Meteorol. 149, 1254–1265.
Grassini, P., van Bussel, L.G.J., van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., van Ittersum, M.K., Cassman, K.G., 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Res. 177, 49–63.
Lobell, D.B., Roberts, J.M., Schlenker, W., Braun, N., Little, B.B., Rejescus, R.M., Hammer, G.L., 2014. Greater sensitivity to drought accompanies maize yield increase in the U.S. midwest. Science 344, 516–519, http://dx.doi.org/10.1126/science.1251423.
Miner, G.L., Hansen, N.C., Inman, D., Sherrod, L.A., Peterson, G.A., 2013. Constraints of no-till dryland agroecosystems as bioenergy production systems. Agron. J. 105, 364–375, http://dx.doi.org/10.2134/agronj2012.0243.
Morell, F.J., Yang, H.S., Cassman, K.G., van Wart, J., Elmore, R.W., Licht, M., Coulter, J.A., Ciampitti, I.A., Pittelkow, C.M., Brouder, S.M., Thomison, P., Lauer, J., Graham, C., Massey, R., Grassini, P., 2016. Can crop simulation models be used to predict local to regional maize yields and total production in the U.S. Corn Belt? Field Crops Res. 192, 1–12, http://dx.doi.org/10.1016/j.fcr.2016.04.004.
Mourtzinis, S., Specht, J.E., Lindsey, L.E., Wiebold, W.J., Ross, J., Nafziger, E.D., Kandel, N., Mueller, H.J., Devillez, P.L., Arriaga, F.J., Conley, S.P., 2015. Climate-induced reduction in US-wide soybean yields underpinned by region- and in-season specific responses. Nat. Plants 1 (2), 14026, http://dx.doi.org/10.1038/nplants.2014.26.

Mourtzinis, S., Ortiz, B.V., Damianidis, D., 2016. Climate change and ENSO effects on southeastern US climate patterns and maize yield. Sci. Rep. 6, 29777, http://dx.doi.org/10.1038/srep29777.
National Aeronautics and Space Administration (NASA). http://power.larc.nasa.gov/.
Overpeck, J.T., Meehl, G.A., Bony, S., Easterling, D.R., 2011. Climate data challenges n the 21st century. Science 331, 700–702.
PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu.
Prein, A.F., Holland, G.J., Rasmussen, R.M., Clark, M.P., Tye, M.R., 2016. Running dry: the U.S. Southwest's drift into a drier climate state. Geophys. Res. Lett. 43, 1272–1279, http://dx.doi.org/10.1002/2015GL066727.
Ramirez-Villegas, J., Challinor, A., 2012. Assessing relevant climate data for agricultural applications. Agric. For. Meteorol. 161, 26–45.
Rezaei, E.E., Siebert, S., Ewert, F., 2015. Impact of data resolution on heat and drought stress simulated for winter wheat in Germany. Europ. J. Agron. 65, 69–82.
Ritchie, S.W., Hanway, J.J., 1982. How a plant crop develops, Spec. Rep. 48. Iowa State Univ. of Sci. and Technol., Coop. Ext. Serv., Ames, IA.
Tang, W.Y., Kassim, A.H.M., Abubakar, S.H., 1996. Comparative studies of various missing data treatment methods Malaysian experience. Atmos. Res. 42, 247–262, http://dx.doi.org/10.1016/0169-8095(95)00067-4.
Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. J. Hydrol. 190, 214–251.
Thornton, P.E., Thornton, M.M., Mayer, B.W., Wilhelmi, N., Wei, Y., Devarakonda, R., Cook, R.B., 2014. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2. ORNL DAAC, Oak Ridge, Tennessee, USA, http://dx.doi.org/10.3334/ORNLDAAC/1219.
Van Ittersum, K.M., Cassman, G.K., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance—a review. Filed Crops Res. 143, 4–17, http://dx.doi.org/10.1016/j.fcr.2012.09.009.
Van Wart, J., Grassini, P., Cassman, K.G., 2013. Impact of derived global weather data on simulated crop yields. Global Change Biol. 19 (13), 3822–3834, http://dx.doi.org/10.1111/gcb.12302.
Van Wart, J., Grassini, P., Yang, H.S., Claessens, L., Jarvis, A., Cassman, K.G., 2015. Creating long-term weather data from the thin air for crop simulation modelling. Agric. For. Meteorol. 208, 49–58, http://dx.doi.org/10.1016/j.agrformet.2015.02.020.
White, J.W., Hoogenboom, G., Wilkens, P.W., Stackhouse, P.W., Hoel, J.M., 2011. Evaluation of satellite-based, modeled derived daily solar radiation data for the continental United States. Agron. J. 103, 1242–1251.
Yang, H.S., Torrion, J., 2013. Weather Interpolator. University of Nebraska (version 2012.1), Available at: http://hybridmaize.unl.edu/weather-interpolator.
Yang, H.S., Dobermann, A., Lindquist, J.L., Walters, D.T., Arkebauer, T.J., Cassman, K.G., 2004. Hybrid-maize—a maize simulation model that combines two crop modeling approaches. Field Crops Res. 87, 131–154.
Yang, H., Dobermann, A., Cassman, K.G., Walters, D.T., 2006. Features, applications, and limitations of the hybrid-maize simulation model. Agron. J. 98, 737–748.
Zhao, G., Siebert, S., Enders, A., Rezaei, E.E., Yan, C., Ewert, F., 2014. Demand for multi-scale weather data for regional crop modeling. Agric. For. Meteorol. 200, 156–171.